

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2005, 18:53:46 ; Search time 163 Seconds
(without alignments)
35.591 Million cell updates/sec

Title: US-09-993-399-1

Perfect score: 79

Sequence: 1 RHYGETKMNQSRSS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	79	100.0	15	6	ABR61469	Human dis
2	79	100.0	355	8	ADG89869	Human kin
3	79	100.0	496	8	ADG89867	Human kin
4	79	100.0	2633	4	ABG06505	Novel hum
5	79	100.0	2863	4	AAW39097	Human pol
6	79	100.0	2663	8	ADQ17932	Human sof
7	79	100.0	2688	4	AAW40883	Human pol
8	69	87.3	2954	2	AAW01632	Amino aci
9	54	68.4	955	2	AAW57365	K39 polyp
10	54	68.4	955	2	AAW03691	Leishmani
11	53	67.1	2013	4	ABW23222	Drosophil
12	52	65.8	154	3	ABW40661	Human ORF
13	52	65.8	154	5	ABP31636	Human str
14	52	65.8	341	5	ABW80079	Human kin
15	52	65.8	341	6	ABG72398	Human kin
16	52	65.8	348	6	ABW44339	Polypepti
17	52	65.8	1279	5	ABW80078	Human kin
18	52	65.8	1279	5	ABG70787	Human kin
19	52	65.8	1279	6	ABG72397	Human par
20	52	65.8	1401	7	ADJ94914	Novel NOV
21	52	65.8	1931	4	ABW61012	Drosophil
22	51	64.6	726	3	AAG31118	Arabidops
23	51	64.6	829	3	AAG31117	Arabidops
24	51	64.6	834	3	AAG31116	Arabidops
25	-50	63.3	677	4	ABW65183	Drosophil

26	49	62.0	172	3	AAW14279	Arabidops
27	49	62.0	346	7	ADC23342	Human kin
28	49	62.0	346	8	ADQ60232	Human mic
29	49	62.0	346	8	ADQ88357	Human mic
30	49	62.0	370	7	ADC23338	Human kin
31	49	62.0	370	8	ADQ60228	Human mic
32	49	62.0	370	8	ADQ88353	Human mic
33	49	62.0	460	3	AAW56650	Human pro
34	49	62.0	487	7	ADC23344	Human kin
35	49	62.0	487	8	ADQ60234	Human mic
36	49	62.0	487	8	ADQ88359	Human mic
37	49	62.0	490	7	ADK40973	Novel hum
38	49	62.0	490	8	ADR15692	Kinase 73
39	49	62.0	512	7	ADC23340	Human kin
40	49	62.0	512	8	ADQ60230	Human mic
41	49	62.0	512	8	ADQ88355	Human mic
42	49	62.0	665	8	ADQ09240	Human KNS
43	49	62.0	665	8	ABW81748	Tumour-as
44	48	60.8	324	2	AAW70235	Leishmani
45	48	60.8	324	5	AAW24949	Leishmani

ALIGNMENTS

RESULT 1

ABR61469
ID ABR61469 standard; peptide; 15 AA.

XX AC ABR61469;

XX DT 01-SEP-2003 (first entry)

XX DE Human disease specific marker CENP-E peptide.

XX KW Human; biopolymer marker; disease state; Alzheimer's disease.

XX OS Homo sapiens.

XX PN WO2003045992-A1.

XX PD 05-JUN-2003.

XX PF 31-OCT-2002; 2002WO-CA001646.

XX PR 23-NOV-2001; 2001US-00993399.

XX PA (SYNX-) SYN.X PHARMA INC.

XX PI Jackowski G, Marshall J;

XX DR WPI; 2003-505186/47.

XX PT New biopolymer marker, useful for indicating, determining risk-assessment of, or identifying therapeutic avenues related to, a disease state e.g., Alzheimer's disease.

XX PS Claim 1; Page 41; 44pp; English.

XX CC The invention relates to a novel biopolymer marker comprising a sequence having 13 amino acids or its analyte, useful in indicating at least one particular disease state. The biopolymer marker is useful for indicating, determining risk-assessment of, or identifying therapeutic avenues related to, a disease state e.g., Alzheimer's disease. The present sequence represents the biopolymer marker of the invention

XX SQ Sequence 15 AA;

Query Match 100.0%; Score 79; DB 6; Length 15;

Best Local Similarity 100.0%; Pred. No. 5.1e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHYGETKMNQSRSS 15

```
Db      |||||
1 RHYGETKNQSSRS 15

RESULT 2
ADG89869
ID ADG89869 standard; protein; 355 AA.
XX
AC ADG89869;
XX
XX
DT 11-MAR-2004 (first entry)
XX
DE Human kinesin motor protein CENP-E340 SEQ ID NO:4.
XX
KW human; CENP-E; HsCENP-E; kinesin motor protein; CENP-E340; cytostatic;
KW neuroprotective; gene therapy; cellular proliferation disorder; cancer;
KW neurological disorder; vesicular transport disorder;
KW hyperproliferative cell growth disorder; centromere-associated protein.
XX
OS Homo sapiens.
XX
PN WO2003104426-A2.
XX
PD 18-DEC-2003.
XX
PF 09-JUN-2003; 2003WO-US018203.
XX
PR 10-JUN-2002; 2002US-0387403P.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Harvey DM, Yang Y, Kohl NE;
XX
DR WPI; 2004-062347/06.
DR N-PSDB; ADG89868.
XX
PT New centromere-associated motor protein, HsCENP-E, useful in diagnosing,
PT treating and preventing disorders characterized by excessive cellular
PT proliferation, including cancer, neurological disorders and disorders of
PT vesicular transport.
XX
PS Disclosure; SEQ ID NO 4; 108pp; English.
XX
CC The present sequence represents a human CENP-E (I) (HsCENP-E) which is a
CC kinesin motor protein, and a centromere-associated protein. The present
CC sequence is more specifically designated CENP-E340. Also described: (1) a
CC composition comprising (I) and a pharmaceutical excipient; (2) a method
CC for screening a compound as an agonist or antagonist of (I); (3) an
CC isolated and purified polynucleotide encoding (I), or which hybridises to
CC (I); (4) a method of detecting polynucleotide; (5) an expression vector
CC comprising the polynucleotide of (3); (6) a host cell comprising the
CC expression vector of (5); (7) a method for producing a polypeptide; (8) a
CC method of modulating cellular proliferation in a mammal; (9) a method of
CC inhibiting HsCENP-E mediated/induced cellular proliferation of a cell in
CC culture; and (10) a method of detecting the presence of cancer in an
CC individual. (I) has cytostatic and neuroprotective activities, and can be
CC used in gene therapy. The HsCENP-E protein can be used in screening
CC assays, in predictive medicine, e.g. diagnostic assays and
CC pharmacogenetics and in treating and preventing disorders characterised
CC by excessive cellular proliferation, including cancer, neurological
CC disorders and disorders of vesicular transport or useful in suppressing
CC hyperproliferative cell growth disorder.
XX
SQ Sequence 355 AA;
Query Match 100.0%; Score 79; DB 8; Length 355;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RHYGETKNQSSRS 15
Db 189 RHYGETKNQSSRS 203

RESULT 3
ADG89867
ID ADG89867 standard; protein; 496 AA.
XX
AC ADG89867;
XX
XX
DT 11-MAR-2004 (first entry)
XX
DE Human kinesin motor protein CENP-E465 SEQ ID NO:2.
XX
KW human; CENP-E; HsCENP-E; kinesin motor protein; CENP-E465; cytostatic;
KW neuroprotective; gene therapy; cellular proliferation disorder; cancer;
KW neurological disorder; vesicular transport disorder;
KW hyperproliferative cell growth disorder; centromere-associated protein.
XX
OS Homo sapiens.
XX
PN WO2003104426-A2.
XX
PD 18-DEC-2003.
XX
PF 09-JUN-2003; 2003WO-US018203.
XX
PR 10-JUN-2002; 2002US-0387403P.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Harvey DM, Yang Y, Kohl NE;
XX
DR WPI; 2004-062347/06.
DR N-PSDB; ADG89866.
XX
PT New centromere-associated motor protein, HsCENP-E, useful in diagnosing,
PT treating and preventing disorders characterized by excessive cellular
PT proliferation, including cancer, neurological disorders and disorders of
PT vesicular transport.
XX
PS Claim 1; SEQ ID NO 2; 108pp; English.
XX
CC The present sequence represents a human CENP-E (I) (HsCENP-E) which is a
CC kinesin motor protein, and a centromere-associated protein. The present
CC sequence is more specifically designated CENP-E465. Also described: (1) a
CC composition comprising (I) and a pharmaceutical excipient; (2) a method
CC for screening a compound as an agonist or antagonist of (I); (3) an
CC isolated and purified polynucleotide encoding (I), or which hybridises to
CC (I); (4) a method of detecting polynucleotide; (5) an expression vector
CC comprising the polynucleotide of (3); (6) a host cell comprising the
CC expression vector of (5); (7) a method for producing a polypeptide; (8) a
CC method of modulating cellular proliferation in a mammal; (9) a method of
CC inhibiting HsCENP-E mediated/induced cellular proliferation of a cell in
CC culture; and (10) a method of detecting the presence of cancer in an
CC individual. (I) has cytostatic and neuroprotective activities, and can be
CC used in gene therapy. The HsCENP-E protein can be used in screening
CC assays, in predictive medicine, e.g. diagnostic assays and
CC pharmacogenetics and in treating and preventing disorders characterised
CC by excessive cellular proliferation, including cancer, neurological
CC disorders and disorders of vesicular transport or useful in suppressing
CC hyperproliferative cell growth disorder.
XX
SQ Sequence 496 AA;
Query Match 100.0%; Score 79; DB 8; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RHYGETKNQSSRS 15
Db 189 RHYGETKNQSSRS 203

RESULT 4
ABG06505
```

ID	ARG06505 standard; protein; 2633 AA.
XX	
AC	ARG06505;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #6496.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Drmanac RT, Liu C, Tang YT;
XX	
DR	WPI: 2001-639362/73.
DR	N-PSDB; AAS70892.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PS	Claim 20; SEQ ID NO 36864; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ARG00010-ARG30377 represent novel human diagnostic
CC	amino acid sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 2633 AA;
	Query Match 100.0%; Score 79; DB 4; Length 2633;
	Best Local Similarity 100.0%; Pred. No. 0.00014;
	Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 RHYGETKNNQRRSSRS 15
Db	189 RHYGETKNNQRRSSRS 203
RESULT 5	
AA03907	
ID	AA03907 standard; protein; 2663 AA.
XX	

AC	AAM39097;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 2242.
XX	
KW	Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US034263.
XX	
PR	23-DEC-1999; 99US-00471275.
PR	21-JAN-2000; 2000US-00488725.
PR	25-APR-2000; 2000US-00552317.
PR	20-JUN-2000; 2000US-00598042.
PR	19-JUL-2000; 2000US-00620312.
PR	03-AUG-2000; 2000US-00653450.
PR	14-SEP-2000; 2000US-00662191.
PR	19-OCT-2000; 2000US-00693036.
PR	29-NOV-2000; 2000US-00727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI	Zhou P, Goodrich R, Drmahac RT;
XX	
DR	WPI; 2001-442253/47.
DR	N-PSDB; AAI58253.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders such
PT	as central nervous system injuries.
XX	
PS	Example 4; SEQ ID NO 2242; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC	encoded polypeptides (AAM38642-AAM42213) with neotropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders. Note: The sequence data for this patent did not form
CC	part of the printed specification
XX	
SQ	Sequence 2663 AA;
	Query Match 100.0%; Score 79; DB 4; Length 2663;
	Best Local Similarity 100.0%; Pred. No. 0.00015;
	Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 RHYGETKMNORSSRS 15
Dd	189 RHYGETKMNORSSRS 203
RESULT	6

ADQ17932
ID ADQ17932 standard; protein; 2663 AA.

XX AC ADQ17932;

XX DT 26-AUG-2004 (first entry)

XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 749.

XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.

XX OS Homo sapiens.

XX PN WO2004048938-A2.

XX PD 10-JUN-2004.

XX PF 26-NOV-2003; 2003WO-US038193.

XX PR 26-NOV-2002; 2002US-0429739P.

XX PA (PROT-) PROTEIN DESIGN LABS INC.

XX PI Aziz N, Ginsburg WM, Zlotnik A;

XX DR WPI; 2004-441208/41.

XX PT Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.

XX PS Example 2; SEQ ID NO 749; 210pp; English.

XX CC The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated protein of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.

XX SQ Sequence 2663 AA;

Query Match 100.0%; Score 79; DB 8; Length 2663;

Best Local Similarity 100.0%; Pred. No. 0.00015; Mismatches 0; Indels 0; Gaps 0;

Matches 15; Conservative 0;

QY 1 RHYGETKNQRRSRS 15

Db 189 RHYGETKNQRRSRS 203

RESULT 7

AAM40883

ID AAM40883 standard; protein; 2688 AA.

XX AC AAM40883;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 5814.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PR 20-JUN-2000; 2000US-00598042.

XX PR 19-JUL-2000; 2000US-00620312.

XX PR 03-AUG-2000; 2000US-00653450.

XX PR 14-SEP-2000; 2000US-00662191.

XX PR 19-OCT-2000; 2000US-00693036.

XX PR 29-NOV-2000; 2000US-00727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Aeundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX PI Zhou P, Goodrich R, Drmanac RT;

XX DR WPI; 2001-442253/47.

XX DR N-PSDB; AAI60039.

XX PT Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

XX PS Example 2; SEQ ID NO 5814; 10078pp; English.

XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification

XX SQ Sequence 2688 AA;

Query Match 100.0%; Score 79; DB 4; Length 2688;

Best Local Similarity 100.0%; Pred. No. 0.00015; Mismatches 0; Indels 0; Gaps 0;

Matches 15; Conservative 0;

QY 1 RHYGETKNQRRSRS 15

Db 213 RHYGETKNQRRSRS 227

RESULT 8

AAY01632

ID AAY01632 standard; protein; 2954 AA.

XX AC AAY01632;

XX DT 22-JUN-1999 (first entry)

XX DE Amino acid sequence of centromere-associated protein-E (CENP-E).

XX KW CENP-E; centromere-associated protein-E; ATPase activity;

KW plus end-directed microtubule motor activity; chromosome congression;
 KW microtubule binding activity; chromosome movement; mitosis;
 KW cell proliferation; tumor; metastasis; vascular malfunction;
 KW inflammatory disease; immune disease; angiogenesis; hypertension;
 KW restenosis; fungal infection; selective herbicide; fungicide;
 KW insecticide; plant growth regulator; activator; cancer cell marker.
 XX
 OS Xenopus sp.
 XX
 XX WO9913061-A1.
 PN
 XX
 XX 18-MAR-1999.
 PD
 XX
 XX 10-SEP-1998; 98WO-US019231.
 PF
 XX
 XX 11-SEP-1997; 97US-0058645P.
 PR
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX
 XX Wood KW, Sakowicz R, Goldstein LSB, Cleveland DW;
 FI
 XX
 XX WPI; 1999-229233/19.
 DR
 XX
 XX N-PSDB; AAX26819.
 DR
 XX
 XX Centromere-associated protein-E and related nucleic acid.
 PT
 XX
 XX
 XX
 XX
 PS Claim 5; Page 66-67; 77pp; English.
 CC
 XX
 CC The present sequence represents CENP-E (centromere-associated protein-E)
 CC of Xenopus. The protein has at least one of plus end-directed microtubule
 CC motor activity, ATPase (adenosine triphosphatase) activity and
 CC microtubule binding activity. CENP-E is the motor that powers chromosome
 CC movement toward microtubule plus ends and is essential for congression of
 CC chromosomes during mitosis. Modulators of CENP-E can thus control cell
 CC proliferation. Agents that modulate CENP-E activity are lead therapeutic,
 CC bioagricultural and diagnostic agents, e.g. for treatment of unwanted
 CC cell proliferation (typical of many examples are tumors and metastases;
 CC vascular malfunction; inflammatory and immune diseases; angiogenesis;
 CC hypertension; restenosis; and fungal infections), also as plant-
 CC protection agents (selective herbicides, fungicides and insecticides) and
 CC plant growth regulators or activators for improving yields. CENP-E is
 CC also a diagnostic marker for dividing cells, including cancer cells
 XX
 XX Sequence 2954 AA;

Query Match 87.3%; Score 69; DB 2; Length 2954;
 Best Local Similarity 86.7%; Pred. No. 0.0099;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RHYGETKNQRRSSRS 15
 |||||
 Db 187 RHYGETKNQRRSSRS 201

RESULT 9
 AAR57365
 ID AAR57365 standard; protein; 955 AA.

XX AAR57365;
 XX
 DT 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 08-MAR-1995 (first entry)
 XX
 XX K39 polypeptide of Leishmania chagasi.

XX Parasite; leishmaniasis; infection; repeat unit; vaccine; antigen;
 KW Leishmania chagasi.
 XX
 OS Leishmania donovani chagasi.

XX WO9416331-A1.
 PN
 XX

PD 21-JUL-1994.
 XX
 XX 10-JAN-1994; 94WO-US000324.
 XX
 PR 15-JAN-1993; 93US-00006676.
 XX
 XX (IASY-) IASYS CORP.
 PA
 XX
 XX Reed SG;
 PI
 XX
 XX WPI; 1994-249402/30.
 DR
 XX
 XX N-PSDB; AAQ70152.
 DR
 XX
 XX Diagnosis of Leishmaniasis - by determining the presence of antibodies
 PT that bind to a K39 repeat unit antigen.
 XX
 XX Disclosure; Page 12-15; 28pp; English.
 PS
 XX
 CC The K39 polypeptide comprises a number of repeated units (described in
 CC AAR57366). Detection of antibodies directed against this repeated unit in
 CC a patients sample is indicative of leishmaniasis. The antigenic repeat
 CC unit can itself be used as a vaccine to protect against infection by a
 CC leishmania parasite. (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 16-OCT-2003 to standardise OS field)
 XX
 XX Sequence 955 AA;

Query Match 68.4%; Score 54; DB 2; Length 955;
 Best Local Similarity 73.3%; Pred. No. 1.4;
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RHYGETKNQRRSSRS 15
 |||||
 Db 241 RHTASTKQNDSSRS 255

RESULT 10
 AAW03691
 ID AAW03691 standard; protein; 955 AA.

XX AAW03691;
 AC
 XX
 DT 16-OCT-2003 (revised)
 DT 09-MAR-1997 (first entry)
 XX

XX Leishmania chagasi K39 antigen.

XX Leishmania chagasi; acidic ribosomal antigen; LcP0; epitope; K39.

XX Leishmania donovani chagasi.
 OS
 XX
 XX WO9633414-A2.
 PN
 XX
 XX 24-OCT-1996.

XX
 XX 19-APR-1996; 96WO-US005472.
 PF
 XX
 XX 21-APR-1995; 95US-00428414.
 PR
 XX
 XX (CORI-) CORIXA CORP.

PA

XX Reed SG;

XX
 XX WPI; 1996-485884/48.
 DR
 XX
 XX N-PSDB; AAT42166.

XX New Leishmania acidic ribosomal P-protein family poly:peptide - used to
 PT develop prods. for diagnosis, detection and protection against Leishmania
 PT infections.

XX Disclosure; Page 36-43; 76pp; English.

PS Compounds including polypeptides that contain at least an epitope of the

CC L. chagasi acidic ribosomal antigen Lcp0 are useful in a variety of
 CC immunoassays for detecting Leishmania infection. Portions of Lcp0
 CC (AAT42164) contg. at least the 17 C-terminal amino acids (AAT42165) have
 CC been found to generate a signal in an ELISA that is equivalent to that
 CC generated by the full length Lcp0. A combination polypeptide may also be
 CC used, comprising an Lcp0 epitope along with an epitope derived from the
 CC Leishmania K39 antigen (AAT42166), pref. the K39 repeat unit antigen
 CC having the sequence given in AA03690. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX
 XX Sequence 955 AA;

Query Match 68.4%; Score 54; DB 2; Length 955;
 Best Local Similarity 73.3%; Pred. No. 1.4;
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 RHYGETKNQRRSSRS 15
 || |||||
 DB 241 RHTASTKNDRSSRS 255

RESULT 11
 ABB62322
 ID ABB62322 standard; protein; 2013 AA.

XX ABB62322;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 13758.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL06425.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.

XX Disclosure; SEQ ID NO 13758; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2013 AA;

Query Match 67.1%; Score 53; DB 4; Length 1013;
 Best Local Similarity 73.3%; Pred. No. 4.6;

Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 OY 1 RHYGETKNQRRSSRS 15
 || |||||
 DB 183 RTVGETNNRRSSRS 197

RESULT 12

AAB40661

ID AAB40661 standard; protein; 154 AA.

XX AAB40661;

XX 08-FEB-2001 (first entry)

XX Human ORFX ORF425 polypeptide sequence SEQ ID NO:850.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antineumatic; antihyroid;
 KW neurodegenerative disorder; cancer; proliferative disorder; hypertension;
 KW antianaemic; gene therapy; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cardiovascular disease; systemic lupus erythematosus; infection;
 KW cholesterol ester storage; severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW severe combined immunodeficiency; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX Homo sapiens.

XX WO2000058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX 31-MAR-1999; 99US-0127607P.

XX 02-APR-1999; 99US-0127636P.

XX 05-APR-1999; 99US-0127728P.

XX 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAC74670.

XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.

XX Claim 11; Page 888-889; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antineumatic; antihyroid; and antianaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester

CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, anti-inflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX
 XX Sequence 154 AA;

Query Match 65.8%; Score 52; DB 3; Length 154;
 Best Local Similarity 66.7%; Pred. No. 0.42;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 RHVGETKNQNRSSRS 15
 DB 29 RHGTGTQNEHSRS 43

RESULT 13
 ABP31636
 ID ABP31636 standard; protein; 154 AA.

AC ABP31636;

DT 09-JUL-2002 (first entry)

DE Human structural protein-like ORF609 protein, SEQ ID NO:1218.

XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.

OS Homo sapiens.

PN WO200190366-A2.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-US017076.

XX 24-MAY-2000; 2000US-0206690P.

XX (CURA-) CURAGEN CORP.

PI Leach MD, Shimkets RA;

DR WPI; 2002-106200/14.

XX N-PSDE; ABN75662.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation.

XX Claim 10; Page 570; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORP (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and

CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antifective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases
 XX

XX Sequence 154 AA;

Query Match 65.8%; Score 52; DB 5; Length 154;
 Best Local Similarity 66.7%; Pred. No. 0.42;

Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 RHVGETKNQNRSSRS 15
 DB 29 RHGTGTQNEHSRS 43

RESULT 14
 ABB80079

ID ABB80079 standard; protein; 341 AA.

AC ABB80079;

DT 27-AUG-2002 (first entry)

DE Human kinesin motor protein (Hskrp5) motor domain amino acid sequence.

XX Human; kinesin motor protein; Hskrp5; cytostatic; cardiovascular;
 KW immunomodulatory; anti-inflammatory; vaccine; gene therapy;
 KW spindle morphogenesis; chromosome; cell division;
 KW cellular proliferation disorder; cancer; hyperplasias; restenosis;
 KW cardiac hypertrophy; immune disorder; inflammation.

OS Homo sapiens.

XX US6379941-B1.

XX 30-APR-2002.

XX 27-NOV-2000; 2000US-00724517.

XX 17-AUG-2000; 2000US-00641807.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C, Freedman R;

XX WPI; 2002-413719/44.

XX Isolated human Kinesin motor protein Hskrp5, useful for preventing,
 PT diagnosing and treating e.g. cancer, hyperplasias, restenosis, cardiac
 PT hypertrophy, immune disorders and inflammation.

XX Claim 4; Fig 4; 29pp; English.

PS The invention relates to an isolated human Kinesin motor protein

CC (HsKrp5). The activity of the protein of the invention may be described

CC as cytoskeletal, cardiovascular, immunomodulatory and anti-inflammatory.

CC The protein may be used in a vaccine or in gene therapy. Kinesin-related

CC proteins (KRPs) participate in spindle morphogenesis and chromosome

CC movement in cell division. The anti-HsKrp5 agonists, antibodies and

CC antagonists may be used to regulate HsKrp5 expression and activity. The

CC anti-HsKrp5 antibodies may also be used as diagnostic agents for

CC detecting the presence of HsKrp5 in samples (e.g. by enzyme linked

CC immunosorbent assay (ELISA)) and therefore diagnose diseases associated

CC with aberrant HsKrp5 expression and activity. They may be used in this

CC way to prevent, diagnose and treat cellular proliferation disorders e.g.

CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders

CC and inflammation. The current sequence represents the human kinesin motor

CC protein (HsKrp5) motor domain amino acid sequence

XX

SQ Sequence 341 AA;

Query Match 65.8%; Score 52; DB 5; Length 341;

Best Local Similarity 66.7%; Pred. No. 1;

Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RHYGETKNQRSSRS 15

|||:|:|:|

Db 192 RHTGTTQNEHSSRS 206

RESULT 15

ABG72398

ID ABG72398 standard; protein; 341 AA.

XX AC ABG72398;

XX DT 11-FEB-2003 (first entry)

XX DE Human kinesin protein, HsKrp5, motor domain.

XX KW Human; kinesin; HsKrp5; ATPase; microtubule; cellular proliferation;

XX KW cancer; hyperplasia; restenosis; cardiac hypertrophy; autoimmune disease;

XX KW immune disorder; arthritis; graft rejection; inflammatory bowel disease;

XX KW inflammation; neurological disorder; vesicular transport disorder;

XX KW enzyme; motor domain.

XX OS Homo sapiens.

XX PN US6448026-B1.

XX PD 10-SEP-2002.

XX PF 27-NOV-2000; 2000US-00723096.

XX PR 17-AUG-2000; 2000US-00641807.

XX PA (CYTO-) CYTOKINETICS INC.

XX PI Beraud C, Freedman R;

XX DR WPI; 2003-089119/08.

XX DR N-PSDB; ABX13614.

XX Screening for modulators of human kinesin protein HsKrp5, which are

PT useful in treating cancers or restenosis, comprises detecting binding or

PT ATPase activity levels of the protein in a first and second concentration

PT of a candidate agent.

XX Claim 8; Fig 4; 30pp; English.

XX The invention relates to screening for modulators of a target protein,

CC comprises detecting the level of binding activity or ATPase activity of

CC the target protein when contacted with a first and second concentration

CC of a candidate agent. The target protein comprises a sequence that has

CC greater than 90% amino acid identity with a sequence of human kinesin

CC protein HsKrp5 (or its motor domain). The method is useful for screening

CC for modulators of a target protein having microtubule stimulated ATPase

CC activity, particularly the human kinesin protein HsKrp5. The modulators

CC of HsKrp5 are useful in diagnosing, preventing or treating cellular

CC proliferation (e.g. cancers (many examples given in the specification) or

CC hyperplasia), restenosis, cardiac hypertrophy, autoimmune disease, immune

CC disorders, arthritis, graft rejection, inflammatory bowel disease,

CC inflammation, neurological disorders, or disorders of vesicular

CC transport. The present sequence represents human HsKrp5, motor domain

XX

SQ Sequence 341 AA;

Query Match 65.8%; Score 52; DB 6; Length 341;

Best Local Similarity 66.7%; Pred. No. 1;

Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RHYGETKNQRSSRS 15

|||:|:|:|

Db 192 RHTGTTQNEHSSRS 206

Search completed: November 1, 2005, 19:10:02

Job time : 166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2005, 18:56:06 ; Search time 39 Seconds
(without alignments)
37.006 Million cell updates/sec

Title: US-09-993-399-1

Perfect score: 79

Sequence: 1 RHYGETKMNQSRSS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	2663	1 S28261	centromere protein
2	69	87.3	2954	2 T14156	kinesin-related pr
3	58	73.4	823	2 T52425	kinesin-like prote
4	58	73.4	888	2 D96619	protein T30B16.9 [
5	54	68.4	955	2 A47334	Lckin kinesin-rela
6	51	64.6	581	2 F84599	probable kinesin h
7	51	64.6	834	2 T06055	hypothetical prote
8	50	63.3	1459	2 T30196	kinesin motor prot
9	49	62.0	198	2 E86183	hypothetical prote
10	49	62.0	665	2 S62328	kinesin-like DNA b
11	48	60.8	1070	2 T06733	kinesin homolog F2
12	47	59.5	670	2 T29898	kinesin protein OS
13	47	59.5	672	2 S54351	kinesin osm-3 - Ca
14	47	59.5	754	2 S48020	kinesin-related pr
15	47	59.5	733	2 S34830	kinesin-related pr
16	46	58.2	102	2 E85814	hypothetical prote
17	46	58.2	108	2 E90966	hypothetical prote
18	46	58.2	127	2 G64960	probable membrane
19	46	58.2	932	2 T49235	kinesin-like prote
20	45	57.0	143	2 C44259	kinesin heavy chai
21	45	57.0	315	2 AF1393	glycosyl transfera
22	45	57.0	315	2 A11768	kinesin-like prote
23	45	57.0	777	2 C85065	kinesin heavy chai
24	45	57.0	963	1 A41919	kinesin heavy chai
25	45	57.0	975	1 A35075	kinesin heavy chai
26	45	57.0	967	1 A31497	kinesin heavy chai
27	45	57.0	1027	2 S37711	kinesin heavy chai
28	45	57.0	1031	1 A38713	kinesin heavy chai
29	45	57.0	1032	2 I38510	neuronal kinesin h

30 44 55.7 294 2 S38983 kinesin-related pr
31 44 55.7 706 1 C42640 kinesin-related pr
32 44 55.7 742 1 S58691 kinesin-related pr
33 44 55.7 987 2 B96766 protein kinesin F2
34 44 55.7 1130 2 T21134 hypothetical prote
35 43 54.4 332 2 C48835 kinesin-like prote
36 43 54.4 701 1 B44259 kinesin-related pr
37 43 54.4 747 1 A57107 kinesin-related pr
38 43 54.4 786 2 A53929 kinesin homolog KH
39 43 54.4 968 2 T45746 hypothetical prote
40 43 54.4 1056 2 C96661 kinesin-like prote
41 43 54.4 1263 2 T13465 hypothetical prote
42 42 53.2 699 1 S38982 kinesin-related pr
43 42 53.2 861 2 T00434 probable kinesin h
44 41 51.9 573 2 F72691 probable fumarate
45 41 51.9 770 1 A44337 kinesin-related pr

ALIGNMENTS

RESULT 1

S28261

centromere protein E - human

N;Alternate names: Centromere

C;Species: Homo sapiens (man)

C;Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004

C;Accession: S28261

R;Yen, T.J.; Li, G.; Schaar, B.T.; Szilak, I.; Cleveland, D.W.

Nature 359, 536-539, 1992

A;Title: CENP-E is a putative kinetochore motor that accumulates just before mitosis.

A;Reference number: S28261; MUID:93024922; PMID:1406971

A;Accession: S28261

A;Molecule type: mRNA

A;Residues: 1-2663 <YEN>

A;Cross-references: UNIPROT:Q02224; EMBL:Z15005; NID:G29864; PIDN:CAA78727.1; PID:G298

C;Genetics:

A;Gene: GDB:CENPE

A;Cross-references: GDB:361164; OMIM:117143

A;Map position: 4q24-q25

C;Superfamily: centromere protein E; kinesin motor domain homology

C;Keywords: ATP; coiled coil; microtubule binding; mitosis; nucleotide binding; P-loop

F;7-335/Domain: kinesin motor domain homology <KMOT>

F;86-93/Region: nucleotide-binding motif A (P-loop)

F;486-2183/Domain: coiled coil #status predicted <COI>

F;92/Binding site: ATP (Lys) #status predicted

Query Match 100.0%; Score 79; DB 1; Length 2663;

Best Local Similarity 100.0%; Pred. No. 1e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RHYGETKMNQSRSS 15

Db 189 RHYGETKMNQSRSS 203

RESULT 2

T14156

kinesin-related protein - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: T14156

R;Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.

Cell 91, 357-366, 1997

A;Title: CENP-E is a plus end-directed kinetochore motor required for metaphase chromo:

A;Reference number: Z17893; MUID:98028574; PMID:9363944

A;Accession: T14156

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-2954 <WOO>

A;Cross-references: UNIPROT:O42263; EMBL:AF027728; NID:G2586070; PID:G2586071; PIDN:AA

C;Genetics:

A;Gene: XCENP-E

C;Superfamily: centromere protein E; kinesin motor domain homology

Query Match 87.3%; Score 69; DB 2; Length 2954;
Best Local Similarity 86.7%; Pred. No. 0.00076;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RHYGETKNQNRSSRS 15
|||||
Db 187 RHYGETKNQNRSSRS 201

RESULT 3

T52425
kinesin-like protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52425
R;Kato, A.; Suzuki, M.; Kuwahara, A.; Ooe, H.; Higano-Inaba, K.; Kameda, Y.
Gene 239, 309-316, 1999
A;Title: Isolation and analysis of cDNA within a 300 kb Arabidopsis thaliana genomic region
A;Reference number: Z25171
A;Accession: T52425
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-823 <KAT>
A;Cross-references: UNIPROT:Q9S7P3; EMBL:AB028468; PIDN:BAA88112.1
C;Genetics:
A;Gene: ZCF125

Query Match 73.4%; Score 58; DB 2; Length 823;
Best Local Similarity 73.3%; Pred. No. 0.021;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RHYGETKNQNRSSRS 15
|||||
Db 184 RHFGETNNVHSSRS 198

RESULT 4

D96619
protein T30E16.9 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D96619
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96619
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-888 <STO>
A;Cross-references: UNIPROT:Q9L062; GB:AE005173; NID:98778739; PIDN:AAF79747.1; GSPDB:GN
C;Genetics:
A;Gene: T30E16.9
A;Map position: 1

Query Match 73.4%; Score 58; DB 2; Length 888;
Best Local Similarity 73.3%; Pred. No. 0.023; 3; Indels 0; Gaps 0;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RHYGETKNQNRSSRS 15
|||||
Db 215 RHFGETNNVHSSRS 229

RESULT 5

A47334
Lckin kinesin-related antigen - Leishmania chagasi (fragment)
C;Species: Leishmania chagasi
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A47334
R;Burns Jr., J.M.; Shreffler, W.G.; Benson, D.R.; Ghali, H.W.; Badaro, R.; Reed, S.G.
Proc. Natl. Acad. Sci. U.S.A. 90, 775-779, 1993
A;Title: Molecular characterization of a kinesin-related antigen of Leishmania chagasi
A;Reference number: A47334; MUID:93133867; PMID:8421715
A;Accession: A47334
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-955 <BUR>
A;Cross-references: UNIPROT:P46865; GB:L07879; NID:G308884; PIDN:AAA29254.1; PID:G3088
A;Experimental source: MHOM/BR/82/BA-2, C1
A;Note: sequence extracted from NCBI backbone (NCBIN:122864, NCBIIP:122865)
C;Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C;Keywords: ATP; nucleotide binding; P-loop
F;13-398/Domain: kinesin motor domain homology <KMOT>
F;122-129/Region: nucleotide-binding motif A (P-loop)

Query Match 68.4%; Score 54; DB 2; Length 955;
Best Local Similarity 73.3%; Pred. No. 0.13; 4; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RHYGETKNQNRSSRS 15
|||||
Db 241 RHTASTKQNRSSRS 255

RESULT 6

F84599
probable kinesin heavy chain [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84599
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84599
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-581 <STO>
A;Cross-references: UNIPROT:Q9SJU7; GB:AE002093; NID:94567265; PIDN:AAD23678.1; GSPDB:
C;Genetics:
A;Gene: At2g21300
A;Map position: 2

Query Match 64.6%; Score 51; DB 2; Length 581;
Best Local Similarity 66.7%; Pred. No. 0.28;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RHYGETKNQNRSSRS 15
|||||
Db 205 RKIGETSLNERSRS 219

RESULT 7

T06055
hypothetical protein F19H22.50 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06055
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15184
A;Accession: T06055
A;Molecule type: DNA
A;Residues: 1-834 <BEV>

A;Cross-references: UNIPROT:Q9SVJ8; EMBL:AL035679; GSPDB:GN00062; ATSP:F19H22.50
A;Experimental source: cultivar Columbia; BAC clone F19H22

C;Genetics:
A;Gene: ATSP:F19H22.50

A;Map position: 4

A;Introns: 63/1; 93/1; 126/3; 164/3; 196/2; 234/3; 276/2; 389/3; 422/3; 633/3; 726/2
C;Superfamily: kinesin motor chain; kinesin motor domain homology
F;14-339/Domain: kinesin motor domain homology <KMOT>

Query Match 64.6%; Score 51; DB 2; Length 834;

Best Local Similarity 66.7%; Pred. No. 0.41; 3; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 0;

Qy 1 RHYGETKMNQSSRS 15

Db 214 RKIGETSLNERSRS 228

RESULT 8

T30196

kinesin motor protein 1 - smut fungus (Ustilago maydis)

C;Species: Ustilago maydis (corn smut)

C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C;Accession: T30196

R;Lehmier, C.; Steinberg, G.; Snetseelaar, K.M.; Schliwa, M.; Kahmann, R.; Bolker, M.

EMBO J. 16, 3464-3473, 1997

A;Title: Identification of a motor protein required for filamentous growth in Ustilago m

A;Reference number: 220770; MUID:97361828; PMID:9218789

A;Accession: T30196

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1459 <LEH>

A;Cross-references: UNIPROT:P87198; EMBL:U92844; NID:g2062749; PID:g2062750; PIDN:AAB633

C;Genetics:

A;Gene: kinl

C;Function:

A;Description: required for filamentous growth in Ustilago maydis

Query Match 63.3%; Score 50; DB 2; Length 1459;

Best Local Similarity 66.7%; Pred. No. 1.1; 4; Indels 0; Gaps 0;

Matches 10; Conservative 1; Mismatches 0;

Qy 1 RHYGETKMNQSSRS 15

Db 487 RHVGATDWNERSRS 501

RESULT 9

E86183

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: E86183

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: E86183

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-198 <STO>

A;Cross-references: UNIPROT:Q9MAU7; GB:AE005172; NID:g7211991; PIDN:AAP40462.1; GSPDB:GN

C;Genetics:

A;Map position: 1

Query Match 62.0%; Score 49; DB 2; Length 198;

Best Local Similarity 72.7%; Pred. No. 0.22; 1; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 0;

Qy 1 RHYGETKMNQ 11

Db 73 RHYNETRINOR 83

RESULT 10

S62328

kinesin-like DNA binding protein KID - human

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 16-Aug-2004

C;Accession: S62328

R;Tokai, N.; Fujimoto-Nishiyama, A.; Toyoshima, Y.; Yonemura, S.; Tsukita, S.; Inoue,

EMBO J. 15, 457-467, 1996

A;Title: Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes an

A;Reference number: S62328; MUID:96174806; PMID:8599929

A;Accession: S62328

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-665 <TOK>

C;Superfamily: kinesin motor domain homology

C;Keywords: ATP; nucleotide binding; P-loop

F;44-374/Domain: kinesin motor domain homology <KMOT>

F;128-135/Region: nucleotide-binding motif A (P-loop)

Query Match 62.0%; Score 49; DB 2; Length 665;

Best Local Similarity 66.7%; Pred. No. 0.76; 3; Indels 0; Gaps 0;

Matches 10; Conservative 2; Mismatches 0;

Qy 1 RHYGETKMNQSSRS 15

Db 232 RTVGATRLNQSSRS 246

RESULT 11

T06733

kinesin homolog F28P10.150 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T06733

R;Quetier, F.; Choisine, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Arti

submitted to the Protein Sequence Database, April 1999

A;Reference number: Z15793

A;Accession: T06733

A;Molecule type: DNA

A;Residues: 1-1070 <QUE>

A;Cross-references: UNIPROT:Q9SV36; EMBL:AL049655; GSPDB:GN00061; ATSP:F28P10.150

A;Experimental source: cultivar Columbia; BAC clone F28P10

C;Genetics:

A;Gene: ATSP:F28P10.150

A;Map position: 3

A;Introns: 113/2; 153/3; 194/3; 256/3; 330/3; 377/1; 427/1; 456/3; 527/3; 572/3; 632/3

C;Superfamily: kinesin-related protein KLP61F; kinesin motor domain homology

F;124-466/Domain: kinesin motor domain homology <KMOT>

Query Match 60.8%; Score 48; DB 2; Length 1070;

Best Local Similarity 66.7%; Pred. No. 1.9; 5; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0;

Qy 1 RHYGETKMNQSSRS 15

Db 311 RHAANTKMNTESSRS 325

RESULT 12

T29898

kinesin protein OSM-3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2004

C;Accession: T29898

R;Nelson, J.; Wohlmann, P.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2005, 18:54:26 ; Search time 171 Seconds
(without alignments)
44.919 Million cell updates/sec

Title: US-09-993-399-1

Perfect score: 79

Sequence: 1 RHYGETXNQRRSRS 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	2663	1 CENE HUMAN	Q02224 homo sapien
2	71	89.9	160	2 O35059	O35059 mus musculus
3	71	89.9	549	2 O7TPX4	O7TPX4 mus musculus
4	71	89.9	2474	2 O6RT24	O6RT24 mus musculus
5	69	87.3	2954	2 O42263	O42263 xenopus lae
6	58	73.4	807	2 Q94RV9	Q94RV9 arabidopsis
7	58	73.4	823	2 Q9S7F3	Q9S7F3 arabidopsis
8	58	73.4	888	2 Q9LQ62	Q9LQ62 arabidopsis
9	54	68.4	859	2 O6IW22	O6IW22 leishmania
10	54	68.4	890	2 O6IW21	O6IW21 leishmania
11	54	68.4	955	1 KINL LEICH	P46865 leishmania
12	53	67.1	2013	2 Q9VKI0	Q9VKI0 drosophila
13	53	67.1	2244	2 Q9NCG0	Q9NCG0 drosophila
14	52	65.8	168	2 O54722	O54722 rattus norv
15	52	65.8	459	2 Q9SS30	Q9SS30 arabidopsis
16	52	65.8	865	2 Q9SL11	Q9SL11 macaca fasc
17	52	65.8	1266	2 O7PCK6	O7PCK6 macaca fasc
18	52	65.8	1304	2 Q86VH0	Q86VH0 homo sapien
19	52	65.8	1335	2 Q86VH1	Q86VH1 homo sapien
20	52	65.8	1394	2 Q7MGZ4	Q7MGZ4 mus musculus
21	52	65.8	1394	2 Q7MGZ5	Q7MGZ5 rattus norv
22	52	65.8	1401	2 Q86VH2	Q86VH2 homo sapien
23	52	65.8	1931	2 Q9NCF9	Q9NCF9 drosophila
24	52	65.8	1931	2 Q9VKH9	Q9VKH9 drosophila
25	51	64.6	581	2 Q9SVJ7	Q9SVJ7 arabidopsis
26	51	64.6	834	2 Q9SVJ8	Q9SVJ8 arabidopsis
27	50	63.3	677	2 Q9VRK9	Q9VRK9 drosophila
28	50	63.3	1459	2 P87198	P87198 ustilago ma
29	50	63.3	1624	2 Q9U679	Q9U679 strongyloc
30	49	62.0	146	2 Q91113	Q91113 morone saxa
31	49	62.0	148	2 O35232	O35232 mus musculus

32	49	62.0	156	2	O15719	O15719 dictyosteli
33	49	62.0	193	2	O84JL8	O84JL8 arabidopsis
34	49	62.0	198	2	O9MAU7	O9MAU7 arabidopsis
35	49	62.0	660	2	O99LC7	O99LC7 mus musculu
36	49	62.0	665	1	KF22_HUMAN	KF22_HUMAN
37	49	62.0	1885	2	O869B8	O869B8 dictyosteli
38	48	60.8	1770	2	O9AVD7	O9AVD7 dictyosteli
39	48	60.8	862	2	O651Z7	O651Z7 oryza sativ
40	48	60.8	925	2	O7YVJ8	O7YVJ8 cryptospori
41	48	60.8	1070	2	O9SVJ6	O9SVJ6 arabidopsis
42	48	60.8	1099	2	O9UI42	O9UI42 leishmania
43	48	60.8	1128	2	O9UI79	O9UI79 leishmania
44	48	60.8	1254	2	O9NEB6	O9NEB6 leishmania
45	47	59.5	347	2	O93XF8	O93XF8 zea mays (m

ALIGNMENTS

RESULT 1
CENE_HUMAN
ID CENE_HUMAN STANDARD; PRT; 2663 AA.
AC Q02224;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Centromeric protein E (CENP-E protein).
GN Name-CENPE;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93024922; PubMed=1406971; DOI=10.1038/359536a0;
RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.;
RT "CENP-E is a putative kinetochore motor that accumulates just before
RT mitosis.";
RL Nature 359:536-539(1992).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=95196755; PubMed=7889940;
RA Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.;
RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed
RT microtubule motor";
RL EMBO J. 14:918-926(1995).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98437347; PubMed=9763420; DOI=10.1083/jcb.143.1.49;
RA Chan G.K.T., Schaar B.T., Yen T.J.;
RT "Characterization of the kinetochore binding domain of CENP-E reveals
RT interactions with the kinetochore proteins CENP-F and hBUBR1.";
RL J. Cell Biol. 143:49-63(1998).
RN [4]
RP FARNESYLATION.
RX MEDLINE=20459117; PubMed=10852915; DOI=10.1074/jbc.M00346200;
RA Ashar H.R., James L., Gray K., Carr D., Black S., Armstrong L.,
RA Bishop W.R., Kirschmeier P.;
RT "Farnesyl transferase inhibitors block the farnesylation of CENP-E and
RT CENP-F and alter the association of CENP-E with the microtubules.";
RL J. Biol. Chem. 275:30451-30457(2000).
CC -!- FUNCTION: Minus-end directed microtubule motor. Probable
CC kinetochore motor. Accumulates just before mitosis at the G2 phase
CC of the cell cycle. Probably important for chromosome movement
CC and/or spindle elongation.
CC -!- SUBUNIT: Interacts with CENP-F and BUBR1 kinase.
CC -!- SUBCELLULAR LOCATION: Associates with kinetochores during
CC congression, relocates to the spindle midzone at anaphase, and is
CC quantitatively discarded at the end of the cell division.
CC -!- SIMILARITY: Belongs to the kinesin-like protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

```

-----
DR EMBL; Z15005; CAA78727.1; -.
DR PIR; S28261; S28261.
DR HSP; P17119; 1F9T.
DR MIM; 117143; -.
DR GENB; HGNC:1856; CENPE.
DR GO; GO:0000776; C:kinetochore; TAS.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003777; F:microtubule motor activity; TAS.
DR GO; GO:0000067; P:DNA replication and chromosome cycle; TAS.
DR GO; GO:0007079; P:mitotic chromosome movement; TAS.
DR GO; GO:0007080; P:mitotic metaphase plate congression; TAS.
DR InterPro; IPR001752; Kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Cell cycle; Cell division; Centromere; Coiled coil;
KW Lipoprotein; Microtubule; Mitosis; Motor protein; Prenylation.
FT DOMAIN 1 335
FT Kinesin-motor.
FT DOMAIN 336 2471
FT Coiled coil (Potential).
FT DOMAIN 2472 2663
FT Globular (Potential).
FT NP_BIND 86 93
FT ATP (By similarity).
FT LIPID 2660 2660
FT S-farnesyl cysteine.
SQ SEQUENCE 2663 AA; 312087 MW; CEFCL3880C8C8CB8 CRC64;
Query Match 100.0%; Score 79; DB 1; Length 2663;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHYGETKNQSRSSRS 15
DB 189 RHYGETKNQSRSSRS 203
|||||
RESULT 2
O35059 PRELIMINARY; PRT; 160 AA.
AC O35059;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Motor domain of KIF10 (Fragment).
GN Name=Cenpe;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Egg;
RC MEDLINE=97420736; PubMed=9275178; DOI=10.1073/pnas.94.18.9654;
RX Nakagawa T., Tanaka Y., Matsuka E., Kondo S., Okada Y., Noda Y.,
RA Kanai Y., Hirokawa N.;
RT "Identification and classification of 16 new kinesin superfamily (KIF)
RL proteins in mouse genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:9654-9659(1997).
DR EMBL; AB001426; BAA22386.1; -.
DR HSP; P17119; 1F9V.
DR MGD; MGI:1098230; Cenpe.
DR GO; GO:0008608; P:spindle kinetochore attachment; TAS.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.

```

```

FT NON_TER 1 1
FT NON_TER 160 160
SQ SEQUENCE 160 AA; 18406 MW; 9E6E4F6E2642C241 CRC64;
Query Match 89.9%; Score 71; DB 2; Length 160;
Best Local Similarity 93.3%; Pred. No. 4.9e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RHYGETKNQSRSSRS 15
DB 108 RHYGETKNQSRSSRS 122
|||||
RESULT 3
Q7TPX4 PRELIMINARY; PRT; 549 AA.
AC Q7TPX4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cenpe protein (Fragment).
GN Name=Cenpe;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Egg;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Egg;
RC Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052843; AAH52843.1; -.
DR HSP; P17119; 1F9T.
DR GO; GO:0008608; P:spindle kinetochore attachment; TAS.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
FT NON_TER 549 549
SQ SEQUENCE 549 AA; 62899 MW; B707C97DC9FB38D3 CRC64;
Query Match 89.9%; Score 71; DB 2; Length 549;
Best Local Similarity 93.3%; Pred. No. 0.0003;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RHYGETKNQSRSSRS 15
DB 108 RHYGETKNQSRSSRS 122
|||||

```

Db 189 RHYGITKNQSRSS 203

RESULT 4

Q6RT24 PRELIMINARY; PRT; 2474 AA.
 ID O6RT24
 AC O6RT24
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Centromere associated protein-E.
 GN Name=Centp;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22806743; PubMed=12925705; DOI=10.1083/jcb.200303167;
 RA Weaver B.A., Bonday Z.Q., Putkey F.R., Kops G.J., Silk A.D.,
 Cleveland D.W.;
 RT "Centromere-associated protein-E is essential for the mammalian
 RT mitotic checkpoint to prevent aneuploidy due to single chromosome
 RT loss.";
 RL J. Cell Biol. 162:551-563 (2003).
 DR EMBL; AY493378; AAR5499.1; -.
 DR GO; GO:0008608; P:spindle kinetochore attachment; TAS.
 DR InterPro; IPR001752; kinesin_motor.
 DR InterPro; IPR002345; Lipocalin.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN 1.
 DR ATP-binding; Microtubule; Motor protein.
 KW ATP-binding; Microtubule; Motor protein.
 SQ SEQUENCE 2474 AA; 286523 MW; BAF52DD0608A2903 CRC64;

Query Match 89.9%; Score 71; DB 2; Length 2474;
 Best Local Similarity 93.3%; Pred. No. 0.0012;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RHYGETKNQSRSS 15

Db 189 RHYGITKNQSRSS 203

RESULT 5

O42263 PRELIMINARY; PRT; 2954 AA.
 ID O42263
 AC O42263
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Kinesin-related protein.
 GN Name=XENP-E;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98028574; PubMed=9363944; DOI=10.1016/S0092-8674(00)80419-5;
 RA Wood K.W., Sakowicz R., Goldstein L.S., Cleveland D.W.;
 RT "XENP-E is a plus end-directed kinetochore motor required for
 RT metaphase chromosome alignment.";
 RL Cell 91:357-366 (1997).
 DR EMBL; AF027728; AAC60300.1; -.
 DR PIR; T14156; T14156.
 DR HSP; P17119; 1P9T.
 DR GO; GO:0005875; C:microtubule associated complex; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0003774; F:motor activity; IEA.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 SQ SEQUENCE 2954 AA; 339966 MW; 439804ED0592679 CRC64;

Query Match 87.3%; Score 69; DB 2; Length 2954;
 Best Local Similarity 86.7%; Pred. No. 0.0033;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RHYGETKNQSRSS 15

Db 187 RHYGETKNQSRSS 201

RESULT 6

O94HV9 PRELIMINARY; PRT; 807 AA.
 ID O94HV9
 AC O94HV9
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Kinesin motor protein (Kin2), putative.
 GN Name=T4M14.11;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Town C.D., Haas B.J., Wu D., Maiti R., Hannick L.I., Chan A.P.,
 Tallon L.J., Rooney T., Utterback T.R., VanAken S.E., Feldblum T.V.,
 White O., Fraser C.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lin X., Kaul S.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Town C.D., Kaul S.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC027036; AAK62792.1; -.
 DR HSP; P20480; IN6M.
 DR GO; GO:0005875; C:microtubule associated complex; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003774; F:motor activity; IEA.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 SQ SEQUENCE 807 AA; 91260 MW; FD7CDAD68EA30C28 CRC64;

Query Match 73.4%; Score 58; DB 2; Length 807;
 Best Local Similarity 73.3%; Pred. No. 0.086;
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 RHYGETKNQSRSS 15

Db 184 RHYGETKNQSRSS 198

RESULT 7

O9S7P3 PRELIMINARY; PRT; 823 AA.
 ID O9S7P3

Q957P3;
 01-MAY-2000 (TrEMBLrel. 13, Created)
 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Kinesin-like protein.
 GN Name=ZCF125;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_taxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ooe H., Kato A., Komeda Y.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20018182; PubMed=10548732; DOI=10.1016/S0378-1119(99)00403-5;
 RA Kato A., Suzuki M., Kuwahara A., Ooe H., Higano-Inaba K., Komeda Y.;
 RT "Isolation and analysis of cDNA within a 300 kb Arabidopsis thaliana
 RT genomic region located around the 100 map unit of chromosome 1.";
 RL Gene 239:309-316(1999).
 DR EMBL; AB028470; BAA88114.1; -;
 DR EMBL; AB028468; BAA88112.1; -;
 DR PIR; T52425; T52425.
 DR HSP; P20480; IN6M.
 DR GO; GO:0005875; C:microtubule associated complex; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003774; F:motor activity; IEA.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS00667; KINESIN_MOTOR_DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 SQ SEQUENCE 823 AA; 93148 MW; 6AFB1C622E4632C9 CRC64;
 Query Match 73.4%; Score 58; DB 2; Length 823;
 Best Local Similarity 73.3%; Pred. No. 0.088;
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 RHYGETKNQSRSSRS 15
 ||:|||||
 DB 184 RHFGTNNVHSSRS 198
 RESULT 8
 Q9LQ62
 ID Q9LQ62 PRELIMINARY; PRT; 888 AA.
 AC Q9LQ62;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE T3OE16.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_taxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Dunn P., Khan S., Kim C., Shinn P.,
 RA Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A., Hansen N.P.,
 RA Huizar L., Kremenetskaia I., Lenz C., Li J., Liu S., Lueros S.,
 RA Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G., Davis R.W.,
 RA Federpiet N.A., Theologis A., Ecker J.R.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bai B., Chin C., Chiou J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,

Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federpiet N.,
 RA Theologis A., Ecker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC009317; AAF79747.1; -;
 DR PIR; D96619; D96619.
 DR HSP; P20480; IN6M.
 DR GO; GO:0005875; C:microtubule associated complex; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003774; F:motor activity; IEA.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS00667; KINESIN_MOTOR_DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 SQ SEQUENCE 888 AA; 100694 MW; 0D640FBACEE01B5 CRC64;
 Query Match 73.4%; Score 58; DB 2; Length 888;
 Best Local Similarity 73.3%; Pred. No. 0.096;
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 RHYGETKNQSRSSRS 15
 ||:|||||
 DB 215 RHFGTNNVHSSRS 229
 RESULT 9
 Q6IW22
 ID Q6IW22 PRELIMINARY; PRT; 859 AA.
 AC Q6IW22;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Kinesin related protein (fragment).
 OS Leishmania donovani.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5661;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KE16;
 RA Sivakumar R., Singh S.;
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY615886; AAT40474.1; -;
 DR GO; GO:0005875; C:microtubule associated complex; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003774; F:motor activity; IEA.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00667; KINESIN_MOTOR_DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 FT NON TER 859 859
 SQ SEQUENCE 859 AA; 95590 MW; 7F941034DACA2155 CRC64;
 Query Match 68.4%; Score 54; DB 2; Length 859;
 Best Local Similarity 73.3%; Pred. No. 0.52;
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RHYGETKNQSRSSRS 15
 ||:|||||
 DB 241 RHTASTKNQSRSSRS 255
 RESULT 10
 Q6IW21
 ID Q6IW21 PRELIMINARY; PRT; 890 AA.
 AC Q6IW21;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)


```

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Kinesin related protein (Fragment).
OS Leishmania donovani.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5661;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DD8;
RA Sivakumar R., Singh S.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY615887; AAT40475.1; -.
DR GO; GO:0005875; C:Microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:Motor activity; IEA.
DR InterPro; IPR001752; F:Motor activity; IEA.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
DR K0138; SUBTILASE_SER; UNKNOWN_1.
KW ATP-binding; Microtubule; Motor protein.
FT NON TER 890
SQ SEQUENCE 890 AA; 98732 MW; 1AFF4AFD0F905407 CRC64;

Query Match 68.4%; Score 54; DB 2; Length 890;
Best Local Similarity 73.3%; Pred. No. 0.54;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RHYGETKMNQSSRS 15
Db 240 RHTASTKMNDRSSRS 254

RESULT 11
KINL_LEICH STANDARD; PRT; 955 AA.
AC P46865;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Kinesin-like protein K39 (Fragment).
GN Name=KIN;
OS Leishmania chagasi.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=44271;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MHOM/BR/92/BA-2;
RX MEDLINE=93133867; PubMed=8421715;
RA Burns J.M. Jr., Shreffler W.G., Benson D.R., Ghalib H.W., Badaro R.,
RA Reed S.G.;
RT "Molecular characterization of a kinesin-related antigen of Leishmania
RT chagasi that detects specific antibody in African and American
RT visceral leishmaniasis."
RT Proc. Natl. Acad. Sci. U.S.A. 90:775-779(1993).
RL -1- DEVELOPMENTAL STAGE: Predominant in amastigotes.
CC -1- SIMILARITY: Belongs to the kinesin-like protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L07879; AAA29254.1; -.
DR PIR; A47334; A47334.
DR HSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.

```

```

DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Coiled coil; Microtubule; Motor protein; Repeat.
FT DOMAIN 1 399 Kinesin-motor (By similarity).
FT DOMAIN 426 >955 Coiled coil (Potential).
FT NP_BIND 122 129 ATP (Potential).
FT DOMAIN 704 >955 7 X 39 AA approximate tandem repeats.
FT REPEAT 704 742 1.
FT REPEAT 743 781 2.
FT REPEAT 782 820 3.
FT REPEAT 821 859 4.
FT REPEAT 860 898 5.
FT REPEAT 899 937 6.
FT REPEAT 938 >955 7 (partial).
FT NON TER 955
SQ SEQUENCE 955 AA; 106168 MW; 8CA76815B8E84C6E9 CRC64;

Query Match 68.4%; Score 54; DB 1; Length 955;
Best Local Similarity 73.3%; Pred. No. 0.58;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RHYGETKMNQSSRS 15
Db 241 RHTASTKMNDRSSRS 255

RESULT 12
Q9VK10 PRELIMINARY; PRT; 2013 AA.
ID Q9VK10;
AC Q9VK10;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG6392-PA.
GN Name=cmet; ORFNames=CG6392;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry J., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muszy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodruff, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang K., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*.";
RT Science 287:2185-2195 (2000).
RL [2]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=2426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards M., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
RT *melanogaster* euchromatic genome sequence.";
RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RL [3]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomics perspective.";
RT Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
RL [4]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RT Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RL [5]
RN
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL [6]
RN
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE003631; AAF53088.2; -;
DR HSSP; P17119; 1F9T.
DR FlyBase; FBgn0040232; cmet.
DR GO; GO:0000776; C:kinetochore of condensed chromosome; NAS.
DR GO; GO:0000940; C:outer kinetochore of condensed chromosome; NAS.
DR GO; GO:0007049; P:cell cycle; IMP.
DR GO; GO:0007080; P:mitotic metaphase plate congression; IMP.
DR InterPro; IPR001752; kinesin_motor.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 2013 AA; 231107 MW; 8A1A465581EF4F5F CRC64;
Query Match 67.1%; Score 53; DB 2; Length 2013;
Best Local Similarity 73.3%; Pred. No. 2.1;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RHYGETNMNRSRS 15
Db 183 RTVGETNMNRSRS 197

RESULT 13

Q9NCGO PRELIMINARY; PRT; 2244 AA.

AC Q9NCGO; 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kinesin-like kinetochore motor protein CENP-meta.
GN Namescmet;
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20351410; PubMed=10893249; DOI=10.1083/jcb.150.1.1;
RA Yucel J.K., Marszalek J.D., McIntosh J.R., Goldstein L.S.B.,
RA Cleveland D.W., Philip A.V.;
RT "CENP-meta, an essential kinetochore kinesis required for the
RT maintenance of metaphase chromosome alignment in *Drosophila*.";
RL J. Cell Biol. 150:1-11(2000).
DR EMBL; AF220353; AAF32355.1; -;
DR HSSP; P17119; 1F9T.
DR FlyBase; FBgn0040232; cmet.
DR GO; GO:0000776; C:kinetochore; IDA.
DR GO; GO:0000940; C:outer kinetochore of condensed chromosome; NAS.
DR GO; GO:0007049; P:cell cycle; IMP.
DR GO; GO:0007080; P:mitotic metaphase plate congression; IMP.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 2244 AA; 257992 MW; F6AA3B2A541ADE0 CRC64;

Query Match 67.1%; Score 53; DB 2; Length 2244;
Best Local Similarity 73.3%; Pred. No. 2.4;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RHYGETNMNRSRS 15
Db 183 RTVGETNMNRSRS 197

RESULT 14

O54722 PRELIMINARY; PRT; 168 AA.

AC O54722; 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kinesin-related protein KRP5 (Fragment).
GN NamesKRP5;
OS *Rattus norvegicus* (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Testes;
RX MEDLINE=96228687; PubMed=8688559;
RA Sperry A.O., Zhao L.P.;
RT "Kinesin-related proteins in the mammalian testes: candidate motors
RT for meiosis and morphogenesis.";

Search completed: November 1, 2005, 19:13:03
Job time : 178 secs

```
RL Mol. Biol. Cell 7:289-305(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testes;
RA Sperry A.O.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035954; AAB88702.1; -.
DR HSSP; P17119; 1F9V.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR001752; kinesin_motor.
DR InterPro; IPR010916; TONB_Box_N.
DR Pfam; PF00225; Kinesin; 1.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW ATP-binding; Microtubule; Motor protein.
FT NON_TER 1
FT NON_TER 168 168
SQ SEQUENCE 168 AA; 18751 MW; D961E4CD1BDAAB7 CRC64;
```

Query Match 65.8%; Score 52; DB 2; Length 168;
Best Local Similarity 66.7%; Pred. No. 0.19;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy 1 RHVGETKNQKRSRS 15
Db 114 RHGTGTTQNNHSSRS 128
||| ||| ||| |||
||| ||| ||| |||

RESULT 15
Q9SS30 Q9SS30 PRELIMINARY; PRT; 459 AA.
AC Q9SS30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative kinesin-like centromere protein.
GN Name=F14P13.22;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X.; Kaul S.; Town C.D.; Benito M.-I.; Creasy T.H.; Haas B.;
RA Renning C.M.; Koo H.; Fujii C.Y.; Utterback T.R.; Barnstead M.E.;
RA Bowman C.L.; White O.; Nierman W.C.; Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009400; AAF02823.1; -.
DR HSSP; P17119; 1F9T.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR001752; kinesin_motor.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 459 AA; 51872 MW; 7C8487E9B7038E6A CRC64;
```

Query Match 65.8%; Score 52; DB 2; Length 459;
Best Local Similarity 73.3%; Pred. No. 0.59;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy 1 RHVGETKNQKRSRS 15
Db 182 RHGTGTTQNNHSSRS 196
||| ||| ||| |||
||| ||| ||| |||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2005, 18:44:05 ; Search time 41 Seconds
(without alignments)
27.311 Million cell updates/sec

Title: US-09-993-399-1

Perfect score: 79

Sequence: 1 RHYGETKNQSRSSRS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/6C COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	2662	4 US-09-595-684B-31	Sequence 31, Appl
2	79	100.0	2662	4 US-09-538-092-1252	Sequence 1252, Ap
3	69	87.3	2954	4 US-09-150-867-1	Sequence 1, Appl
4	54	68.4	955	1 US-08-006-676B-1	Sequence 1, Appl
5	54	68.4	955	1 US-08-282-845-2	Sequence 2, Appl
6	54	68.4	955	2 US-08-428-414A-3	Sequence 3, Appl
7	54	68.4	955	5 PCT-US94-00324-1	Sequence 1, Appl
8	52	65.8	341	3 US-09-724-517-4	Sequence 4, Appl
9	52	65.8	341	4 US-09-641-807A-4	Sequence 4, Appl
10	52	65.8	341	4 US-09-723-096-4	Sequence 4, Appl
11	52	65.8	1279	3 US-09-724-517-2	Sequence 2, Appl
12	52	65.8	1279	4 US-09-641-807A-2	Sequence 2, Appl
13	52	65.8	1279	4 US-09-723-096-2	Sequence 2, Appl
14	49	62.0	346	3 US-09-724-224-6	Sequence 6, Appl
15	49	62.0	346	4 US-10-093-317-6	Sequence 6, Appl
16	49	62.0	370	3 US-09-724-224-2	Sequence 2, Appl
17	49	62.0	370	4 US-10-093-317-2	Sequence 2, Appl
18	49	62.0	487	3 US-09-724-224-8	Sequence 8, Appl
19	49	62.0	487	4 US-10-093-317-8	Sequence 8, Appl
20	49	62.0	490	4 US-09-949-016-8784	Sequence 8784, Ap
21	49	62.0	512	3 US-09-724-224-4	Sequence 4, Appl
22	49	62.0	512	4 US-10-093-317-4	Sequence 4, Appl
23	49	62.0	665	4 US-09-595-684B-35	Sequence 35, Appl
24	48	60.8	324	3 US-09-183-861-53	Sequence 53, Appl
25	48	60.8	324	3 US-09-022-765-53	Sequence 53, Appl
26	48	60.8	324	4 US-09-551-974A-53	Sequence 53, Appl
27	48	60.8	324	4 US-09-565-501A-53	Sequence 53, Appl

Sequence 53, Appl
Sequence 53, Appl
Sequence 4, Appl
Sequence 3, Appl
Sequence 7425, Ap
Sequence 17, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 22, Appl
Sequence 1060, Ap
Sequence 21, Appl
Sequence 19, Appl
Sequence 10196, A
Sequence 27, Appl
Sequence 24, Appl
Sequence 26, Appl
Sequence 1293, Ap
Sequence 26189, A

ALIGNMENTS

RESULT 1
US-09-595-684B-31
; Sequence 31, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; TYPE: PRT
; ORGANISM: Human
US-09-595-684B-31

Query Match 100.0%; Score 79; DB 4; Length 2662;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHYGETKNQSRSSRS 15
Db 189 RHYGETKNQSRSSRS 203

RESULT 2

US-09-538-092-1252
; Sequence 1252, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01

```
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratPatSeqFormatter Version 0.9
; SEQ ID NO 1252
; LENGTH: 2663
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q02224
US-09-538-092-1252

Query Match      100.0%; Score 79; DB 4; Length 2663;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RHYGETKNQRRSSRS 15
Db      189 RHYGETKNQRRSSRS 203

RESULT 3
US-09-150-867-1
; Sequence 1, Application US/09150867
; Patent No. 6645748
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth W.
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S.B.
; APPLICANT: Cleveland, Don W.
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Plus End-Directed Microtubule Motor Required for
; FILE OF INVENTION: Chromosome Congression
; CURRENT FILING DATE: 1998-09-10
; EARLIER FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2954
; TYPE: PRT
; ORGANISM: Xenopus sp.
; FEATURE:
; OTHER INFORMATION: Xenopus centromere-associated protein-E (XCENP-E)
; OTHER INFORMATION: member of the kinesin superfamily of microtubule
; OTHER INFORMATION: motor proteins
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(472)
; OTHER INFORMATION: kinesin like motor domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (473)..(2752)
; OTHER INFORMATION: rod domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (2753)..(2954)
; OTHER INFORMATION: tail domain
US-09-150-867-1

Query Match      87.3%; Score 69; DB 4; Length 2954;
Best Local Similarity 86.7%; Pred. No. 0.0014;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 RHYGETKNQRRSSRS 15
Db      187 RHYGETKNQRRSSRS 201

RESULT 4
US-08-006-676B-1
```

```
; Sequence 1, Application US/08006676B
; Patent No. 5411865
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; TITLE OF INVENTION: Diagnosis of Leishmaniasis
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeffrey B. Oster
; STREET: 8339 SE 57th Street
; CITY: Mercer Island
; STATE: Washington
; COUNTRY: USA
; ZIP: 98040-4906
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/006,676B
; FILING DATE: 15-JAN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oster, Jeffrey B.
; REGISTRATION NUMBER: 32,585
; REFERENCE/DOCKET NUMBER: REED-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 232 7845
; TELEFAX: (206) 236 0205
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 955 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-006-676B-1

Query Match      68.4%; Score 54; DB 1; Length 955;
Best Local Similarity 73.3%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 RHYGETKNQRRSSRS 15
Db      241 RHTASTKMDRSSRS 255

RESULT 5
US-08-282-845-2
; Sequence 2, Application US/08282845
; Patent No. 5719263
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: A 230kd Antigen Present in Leishmania
; TITLE OF INVENTION: Species
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Apple Macintosh
; SOFTWARE: Microsoft Word for Macintosh 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/006,676
```

1 FILING DATE: JANUARY 15, 1993
2 CLASSIFICATION: 435
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Perkins, Patricia Anne
5 REGISTRATION NUMBER: 34,693
6 REFERENCE/DOCKET NUMBER: 5004-A
7 TELECOMMUNICATION INFORMATION:
8 TELEPHONE: (206)587-0430
9 TELEFAX: (206)233-0644
10 INFORMATION FOR SEQ ID NO: 2:
11 SEQUENCE CHARACTERISTICS:
12 LENGTH: 955 amino acids
13 TYPE: amino acid
14 TOPOLOGY: linear
15 MOLECULE TYPE: protein
16 US-08-282-845-2

Query Match 68.4%; Score 54; DB 1; Length 955;
Best Local Similarity 73.3%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RHYGETKMNORSSRS 15
Db 241 RHTASTKMNDRSSRS 255

1 RESULT 6
2 US-08-428-414A-3
3 Sequence 3, Application US/08428414A
4 Patent No. 5912166
5 GENERAL INFORMATION:
6 APPLICANT: Reed, Steven G.
7 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
8 LEISHMANIASIS
9 NUMBER OF SEQUENCES: 5
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: SEED and BERRY
12 STREET: 6300 Columbia Center, 701 Fifth Avenue
13 CITY: Seattle
14 STATE: Washington
15 COUNTRY: USA
16 ZIP: 98104-7092
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 OPERATING SYSTEM: IBM PC compatible
20 SOFTWARE: PatentIn Release #1.0, Version #1.30
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/08/428,414A
23 FILING DATE: 21-APR-1995
24 CLASSIFICATION: 436
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Kadlecsek, Ann T.
27 REGISTRATION NUMBER: 39,244
28 REFERENCE/DOCKET NUMBER: 210121.407
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (206) 622-4900
31 TELEFAX: (206) 682-6031
32 TELEX: 3723836 SEEDANDBERRY
33 INFORMATION FOR SEQ ID NO: 3:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 955 amino acids
36 TYPE: amino acid
37 STRANDEDNESS:
38 TOPOLOGY: linear
39 US-08-428-414A-3

Query Match 68.4%; Score 54; DB 2; Length 955;
Best Local Similarity 73.3%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RHYGETKMNORSSRS 15
Db 241 RHTASTKMNDRSSRS 255

1 Db 241 RHTASTKMNDRSSRS 255
2 PCT-US94-00324-1
3 RESULT 7
4 Sequence 1, Application PC/TUS9400324
5 GENERAL INFORMATION:
6 APPLICANT: Reed, Steven
7 TITLE OF INVENTION: Diagnosis of Leishmaniasis
8 NUMBER OF SEQUENCES: 3
9 CORRESPONDENCE ADDRESS:
10 ADDRESSEE: Immunex Corporation
11 STREET: 51 University Street
12 CITY: Seattle
13 STATE: Washington
14 COUNTRY: USA
15 ZIP: 98101
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: Floppy disk
18 COMPUTER: Apple Macintosh
19 OPERATING SYSTEM: Apple System 7.1
20 SOFTWARE: Microsoft Word, version 5.1a
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: PCT/US94/00324
23 FILING DATE:
24 CLASSIFICATION:
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: US/08/006,676
27 FILING DATE: 15-JAN-1993
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Perkins, Patricia Anne
30 REGISTRATION NUMBER: 34,693
31 REFERENCE/DOCKET NUMBER: 5004-WO
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: (206) 587-0430
34 TELEFAX: (206) 233-0644
35 INFORMATION FOR SEQ ID NO: 1:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 955 amino acids
38 TYPE: amino acid
39 TOPOLOGY: linear
40 MOLECULE TYPE: protein
41 PCT-US94-00324-1

Query Match 68.4%; Score 54; DB 5; Length 955;
Best Local Similarity 73.3%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RHYGETKMNORSSRS 15
Db 241 RHTASTKMNDRSSRS 255

1 RESULT 8
2 US-09-724-517-4
3 Sequence 4, Application US/09724517
4 Patent No. 6379941
5 GENERAL INFORMATION:
6 APPLICANT: Beraud, Christophe
7 APPLICANT: Freedman, Richard
8 TITLE OF INVENTION: No. 6379941el motor proteins and methods for
9 their use
10 FILE REFERENCE: 1031
11 CURRENT APPLICATION NUMBER: US/09/724,517
12 CURRENT FILING DATE: 2000-11-27
13 PRIOR APPLICATION NUMBER: US/09/641,807
14 PRIOR FILING DATE: 2000-08-17
15 NUMBER OF SEQ ID NOS: 4
16 SOFTWARE: FastSeq for Windows Version 4.0
17 SEQ ID NO 4
18 LENGTH: 341
19 TYPE: PRT
20 ORGANISM: Human

US-09-724-517-4

Query Match 65.8%; Score 52; DB 3; Length 341;
Best Local Similarity 66.7%; Pred. No. 0.17;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHYGETKMQRSRS 15
DB 192 RHTGTTQNEHSRS 206

RESULT 9

US-09-641-807A-4
; Sequence 4, Application US/09641807A
; Patent No. 6440731
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6440731 motor proteins and methods for
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/641,807A
; CURRENT FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-641-807A-4

Query Match 65.8%; Score 52; DB 4; Length 341;
Best Local Similarity 66.7%; Pred. No. 0.17;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHYGETKMQRSRS 15
DB 192 RHTGTTQNEHSRS 206

RESULT 10

US-09-723-096-4
; Sequence 4, Application US/09723096
; Patent No. 6448026
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6448026 motor proteins and methods for
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/723,096
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US/09/641,807
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-723-096-4

Query Match 65.8%; Score 52; DB 4; Length 341;
Best Local Similarity 66.7%; Pred. No. 0.17;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHYGETKMQRSRS 15
DB 192 RHTGTTQNEHSRS 206

RESULT 11

US-09-724-517-2

; Sequence 2, Application US/09724517
; Patent No. 6379941

; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6379941 motor proteins and methods for
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/724,517
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US/09/641,807
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Human
; NAME/KEY: VARIANT
; LOCATION: (409)...(436)
; OTHER INFORMATION: Xaa = any amino acid
US-09-724-517-2

Query Match 65.8%; Score 52; DB 3; Length 1279;
Best Local Similarity 66.7%; Pred. No. 0.7;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHYGETKMQRSRS 15
DB 194 RHTGTTQNEHSRS 208

RESULT 12

US-09-641-807A-2
; Sequence 2, Application US/09641807A
; Patent No. 6440731
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6440731 motor proteins and methods for
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/641,807A
; CURRENT FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Human
; NAME/KEY: VARIANT
; LOCATION: (409)...(446)
; OTHER INFORMATION: Xaa = any amino acid
US-09-641-807A-2

Query Match 65.8%; Score 52; DB 4; Length 1279;
Best Local Similarity 66.7%; Pred. No. 0.7;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHYGETKMQRSRS 15
DB 194 RHTGTTQNEHSRS 208

RESULT 13

US-09-723-096-2
; Sequence 2, Application US/09723096
; Patent No. 6448026
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard

; TITLE OF INVENTION: No. 6448026el motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/723,096
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US/09/641,807
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (409)...(436)
; OTHER INFORMATION: Xaa = any amino acid
US-09-723-096-2

Query Match 65.8%; Score 52; DB 4; Length 1279;
Best Local Similarity 66.7%; Pred. No. 0.7;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHYGETKMQRSSRS 15
DB 194 RHTGTTQNEHSSRS 208

RESULT 14
US-09-724-224-6
; Sequence 6, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-6

Query Match 62.0%; Score 49; DB 3; Length 346;
Best Local Similarity 66.7%; Pred. No. 0.61;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHYGETKMQRSSRS 15
DB 208 RTVGATRLNQRSSRS 222

RESULT 15
US-10-093-317-6
; Sequence 6, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-6

Query Match 62.0%; Score 49; DB 4; Length 346;
Best Local Similarity 66.7%; Pred. No. 0.61;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHYGETKMQRSSRS 15
DB 208 RTVGATRLNQRSSRS 222

Search completed: November 1, 2005, 18:54:21
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2005, 18:50:35 ; Search time 166 Seconds
(without alignments)
37.775 Million cell updates/sec

Title: US-09-993-399-1

Perfect score: 79.

Sequence: 1 RHYGETKMNQSRSS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	355	18 US-10-517-510-4	Sequence 4, Appli
2	79	100.0	496	18 US-10-517-510-2	Sequence 2, Appli
3	79	100.0	2503	17 US-10-828-985A-11	Sequence 11, Appl
4	79	100.0	2543	17 US-10-828-985A-9	Sequence 9, Appli
5	79	100.0	2568	17 US-10-828-985A-7	Sequence 7, Appli
6	79	100.0	2633	18 US-10-450-763-36864	Sequence 36864, A
7	79	100.0	2663	16 US-10-723-860-749	Sequence 749, App
8	69	87.3	2954	18 US-10-650-280-1	Sequence 1, Appli
9	65	82.3	668	16 US-10-425-115-263022	Sequence 263022,
10	65	82.3	694	15 US-10-425-114-59725	Sequence 59725, A
11	63	79.7	201	16 US-10-437-963-116808	Sequence 116808,

53	67.1	2013	20	US-11-097-143-13758	Sequence 13758, A
52	65.8	154	11	US-09-864-408A-1218	Sequence 1218, Ap
52	65.8	1382	16	US-10-437-963-176714	Sequence 176714,
52	65.8	1401	15	US-10-287-226-142	Sequence 142, App
52	65.8	1931	20	US-11-097-143-9828	Sequence 9828, Ap
51	64.6	197	16	US-10-425-115-245875	Sequence 245875,
18	63.3	677	20	US-11-097-143-22341	Sequence 22341, A
49	62.0	259	16	US-10-425-115-210540	Sequence 210540,
49	62.0	346	16	US-10-797-893-6	Sequence 6, Appli
49	62.0	370	16	US-10-797-893-2	Sequence 2, Appli
49	62.0	460	9	US-09-925-300-1228	Sequence 1228, Ap
49	62.0	487	16	US-10-797-893-8	Sequence 8, Appli
49	62.0	490	15	US-10-334-143-80	Sequence 80, Appl
49	62.0	512	16	US-10-797-893-4	Sequence 4, Appli
49	62.0	665	16	US-10-733-878-425	Sequence 425, App
48	60.8	324	9	US-09-874-923-53	Sequence 53, Appl
48	60.8	324	9	US-09-991-496-53	Sequence 53, Appl
48	60.8	420	16	US-10-437-963-114373	Sequence 114373,
48	60.8	821	16	US-10-437-963-114370	Sequence 114370,
48	60.8	906	16	US-10-437-963-163527	Sequence 163527,
47	59.5	54	10	US-09-826-734-144	Sequence 144, App
47	59.5	75	15	US-10-424-599-152360	Sequence 152360,
47	59.5	78	16	US-10-425-115-201216	Sequence 201216,
35	59.5	247	15	US-10-424-599-156142	Sequence 156142,
47	59.5	272	15	US-10-424-599-274019	Sequence 274019,
47	59.5	458	15	US-10-425-114-59085	Sequence 59085, A
37	59.5	670	15	US-10-369-493-5967	Sequence 5967, Ap
47	59.5	789	16	US-10-425-115-200354	Sequence 200354,
40	58.2	782	16	US-10-437-963-173873	Sequence 173873,
41	58.2	788	16	US-10-437-963-173884	Sequence 173884,
46	58.2	932	16	US-10-473-276-1	Sequence 1, Appli
43	57.0	114	16	US-10-437-963-113879	Sequence 113879,
45	57.0	252	16	US-10-767-701-36255	Sequence 36255, A
45	57.0	312	15	US-10-424-599-174384	Sequence 174384,

ALIGNMENTS

RESULT 1

US-10-517-510-4
; Sequence 4, Application US/10517510
; Publication No. US20050164201A1
; GENERAL INFORMATION:
; APPLICANT: Harvey, Diane Marie
; APPLICANT: Yang, Yi
; APPLICANT: Kohl, Nancy
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE ENCODING
; TITLE OF INVENTION: A NOVEL CENTROMERE-ASSOCIATED MOTOR PROTEIN, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 21023P
; CURRENT APPLICATION NUMBER: US/10/517,510
; CURRENT FILING DATE: 2004-12-10
; PRIOR APPLICATION NUMBER: PCT/US03/18203
; PRIOR FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: 60/387,403
; PRIOR FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Human
US-10-517-510-4

Query Match 100.0%; Score 79; DB 18; Length 355;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHYGETKMNQSRSS 15

|||||

189 RHYGETKMNQSRSS 203

Db


```
; Sequence 36864, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790C13/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 36864
; LENGTH: 2633
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (77)..(99)
; OTHER INFORMATION: KINESIN HEAVY CHAIN SIGNATURE domain identified by eMATRIX,
; accession number FRO0380A, p-value=1.000e-25, raw score of 14.18
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (12)..(2412)
; OTHER INFORMATION: Kinesin motor domain identified by Pfam, accession name
; OTHER INFORMATION: kinesin, E-value=4.9e-217, Pfam score of 734.4
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2633)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-36864

Query Match 100.0%; Score 79; DB 18; Length 2633;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHYGETKMNQSRSS 15
Db 189 RHYGETKMNQSRSS 203

RESULT 7
US-10-723-860-749
; Sequence 749, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 749
; LENGTH: 2663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-749

Query Match 100.0%; Score 79; DB 16; Length 2663;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHYGETKMNQSRSS 15
```

```
Db 189 RHYGETKMNQSRSS 203

RESULT 8
US-10-650-280-1
; Sequence 1, Application US/10650280
; Publication No. US20050191631A1
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth W.
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S.B.
; APPLICANT: Cleveland, Don W.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Plus End-Directed Microtubule Motor Required for
; TITLE OF INVENTION: Chromosome Congression
; FILE REFERENCE: 18557C-000110US
; CURRENT APPLICATION NUMBER: US/10/650,280
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/150,867
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,645
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2954
; TYPE: PRT
; ORGANISM: Xenopus sp.
; FEATURE:
; OTHER INFORMATION: Xenopus centromere-associated protein-E (XCENP-E)
; OTHER INFORMATION: member of the kinesin superfamily of microtubule
; OTHER INFORMATION: motor proteins
; NAME/KEY: DOMAIN
; LOCATION: (1)..(472)
; OTHER INFORMATION: kinesin like motor domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (473)..(2752)
; OTHER INFORMATION: rod domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (2753)..(2954)
; OTHER INFORMATION: tail domain
US-10-650-280-1

Query Match 87.3%; Score 69; DB 18; Length 2954;
Best Local Similarity 86.7%; Pred. No. 0.0081;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RHYGETKMNQSRSS 15
Db 187 RHYGETKMNQSRSS 201

RESULT 9
US-10-425-115-263022
; Sequence 263022, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 263022
; LENGTH: 668
```

; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_171488C.1.pep
US-10-425-115-263022

Query Match 82.3%; Score 65; DB 16; Length 668;
Best Local Similarity 80.0%; Pred. No. 0.0086;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RHYGETKNQSRSS 15
||:|||||
Db 195 RHFGETNMNRSRS 209

RESULT 10
US-10-425-114-59725
; Sequence 59725, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59725
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LFB3912-011-C4_FLI.pep
US-10-425-114-59725

Query Match 82.3%; Score 65; DB 15; Length 694;
Best Local Similarity 80.0%; Pred. No. 0.0089;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RHYGETKNQSRSS 15
||:|||||
Db 221 RHFGETNMNRSRS 235

RESULT 11
US-10-437-963-116808
; Sequence 116808, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 116808
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_20274C.1.pep
US-10-437-963-116808

Query Match 79.7%; Score 63; DB 16; Length 201;
Best Local Similarity 80.0%; Pred. No. 0.0054;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RHYGETKNQSRSS 15
||:|||||
Db 41 RHFGETNMNRSRS 55

RESULT 12
US-11-097-143-13758
; Sequence 13758, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13758
; LENGTH: 2013
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-13758

Query Match 67.1%; Score 53; DB 20; Length 2013;
Best Local Similarity 73.3%; Pred. No. 3.8;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RHYGETKNQSRSS 15
||:|||||
Db 183 RTVGETNMNRSRS 197

RESULT 13
US-09-864-408A-1218
; Sequence 1218, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides En
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1218
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-09-864-408A-1218

Query Match 65.8%; Score 52; DB 11; Length 154;
Best Local Similarity 66.7%; Pred. No. 0.37;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHYGETKNQNRSSRS 15
||| |.||. ||||
Db 29 RHTGTTQWNEHSSRS 43

RESULT 14

US-10-437-963-176714
; Sequence 176714, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 176714

; LENGTH: 1382
; TYPE: PRT
; ORGANISM: Oryza sativa

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_74437C.1.pep

US-10-437-963-176714

Query Match 65.8%; Score 52; DB 16; Length 1382;
Best Local Similarity 73.3%; Pred. No. 3.9;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RHYGETKNQNRSSRS 15
||| |.||. ||||
Db 180 RHIGETNMVYSSRS 194

RESULT 15

US-10-287-226-142
; Sequence 142, Application US/10287226
; Publication No. US20040086875A1

; GENERAL INFORMATION:

; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsobrook, John P.,
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Ference,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiPippo, Vincent A.,
; APPLICANT: Edinger, Shlomit R.,
; APPLICANT: Eisen, Andrew,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gangolli, Esha A.,
; APPLICANT: Gorman, Linda,

; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khramtsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyankar, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.,
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Suresh G.,
; APPLICANT: Spaderna, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Vernhet, Corine A.M.,
; APPLICANT: Zerhusen, Bryan D.,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 142
; LENGTH: 1401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-226-142

Query Match 65.8%; Score 52; DB 15; Length 1401;
Best Local Similarity 66.7%; Pred. No. 3.9;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHYGETKNQNRSSRS 15
||| |.||. ||||
Db 193 RHTGTTQWNEHSSRS 207

Search completed: November 1, 2005, 19:07:11
Job time : 168 secs